



日本国特許庁

PATENT OFFICE  
JAPANESE GOVERNMENT

08/263125

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

1993年 6月29日

出願番号

Application Number:

平成 5年特許願第182138号

願 人

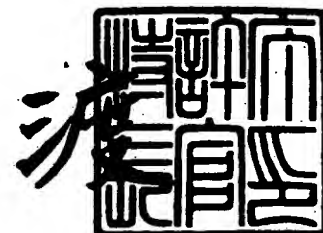
Applicant (s):

ソニー株式会社

1994年 4月 8日

特許庁長官  
Commissioner,  
Patent Office

麻生



【書類名】 特許願

【整理番号】 S93036049

【提出日】 平成 5年 6月29日

【あて先】 特許庁長官 麻生 渡 殿

【国際特許分類】 H04B 15/00

【発明の名称】 音声信号送受信装置

【請求項の数】 13

【発明者】

    【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

    【氏名】 片柳 恵一

【発明者】

    【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

    【氏名】 小高 健太郎

【発明者】

    【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

    【氏名】 西口 正之

【特許出願人】

    【識別番号】 000002185

    【氏名又は名称】 ソニー株式会社

    【代表者】 大賀 典雄

【代理人】

    【識別番号】 100067736

    【弁理士】

    【氏名又は名称】 小池 晃

【代理人】

    【識別番号】 100086335

【弁理士】

【氏名又は名称】 田村 榮一

【代理人】

【識別番号】 100096677

【弁理士】

【氏名又は名称】 伊賀 誠司

【手数料の表示】

【納付方法】 予納

【予納台帳番号】 019530

【納付金額】 14,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9102452

【書類名】 明細書

【発明の名称】 音声信号送受信装置

【特許請求の範囲】

【請求項1】 デジタル信号処理により音声信号を高効率で圧縮する音声信号送信用符号化回路を有する音声信号送受信装置において、

上記音声信号送信用符号化回路で得られる分析パラメータを用いて雑音区間を検出する雑音区間検出手段と、

上記雑音区間検出手段で検出された雑音区間の雑音レベルを検出する雑音レベル検出手段と、

上記雑音レベル検出手段で検出された雑音レベルに応じて受話音量を制御する制御手段とを有することを特徴とする音声信号送受信装置。

【請求項2】 上記雑音区間検出手段は、入力音声信号の1フレームにつき上記分析パラメータとして1次の線形予測符号化係数を用い、該1次の線形予測符号化係数が所定のしきい値よりも小さいときには該1フレームを雑音区間とすることを特徴とする請求項1記載の音声信号送受信装置。

【請求項3】 上記雑音区間検出手段は、入力音声信号の1フレームにつき分析パラメータとしてピッチ成分の強弱を示すピッチゲインを用い、該ピッチゲインが所定の範囲内にあるときに該1フレームを雑音区間とすることを特徴とする請求項1又は2記載の音声信号送受信装置。

【請求項4】 上記雑音区間検出手段は、上記分析パラメータとしてピッチ周波数を用い、入力音声信号の1フレームの該ピッチ周波数成分が0であるときに該1フレームを雑音区間とすることを特徴とする請求項1、2又は3記載の音声信号送受信装置。

【請求項5】 上記雑音区間検出手段は、上記分析パラメータとしてフレームパワーを用い、入力音声信号の1フレームの該フレームパワーが所定のしきい値よりも小さいときに該1フレームを雑音区間とすることを特徴とする請求項1、2、3又は4記載の音声信号送受信装置。

【請求項6】 上記雑音区間検出手段は、現在のフレームと過去のフレームでのフレームパワーの変化量が所定のしきい値を越えたときには、現在のフレーム

を雑音区間としていても、該現在のフレームを音声区間とすることを特徴とする請求項1、2、3、4又は5記載の音声信号送受信装置。

【請求項7】 上記雑音区間検出手段は、複数連続フレームの上記分析パラメータの値を考慮して、雑音区間の検出を行うことを特徴とする請求項2、3又は5記載の音声信号送受信装置。

【請求項8】 上記雑音レベル検出手段は、上記雑音区間検出手段で検出された雑音区間の雑音レベル出力に対してフィルタ処理を施すことを特徴とする請求項1記載の音声信号送受信装置。

【請求項9】 上記雑音レベル検出手段が雑音レベル出力に対して施すフィルタ処理は、最小値フィルタ処理であることを特徴とする請求項8記載の音声信号送受信装置。

【請求項10】 送信部及び受信部を備える音声信号送受信装置において、上記送信部に送話音声入力が無いときに送話用マイクロフォンに入力される音声レベルを雑音レベルとして検出する雑音レベル検出手段と、

上記雑音レベル検出手段で検出された雑音レベルに応じて受話音量を制御する制御手段とを有することを特徴とする音声信号送受信装置。

【請求項11】 上記雑音レベル検出手段は、上記送信部の上記送話用マイクロフォンに送信通話用電源オン直後に入力される音声レベルを検出することを特徴とする請求項10記載の音声信号送受信装置。

【請求項12】 上記雑音レベル検出手段は、上記送信部の着信信号の待機状態において、所定期間毎に上記送信用マイクロフォンに入力される音声レベルを検出することを特徴とする請求項10記載の音声信号送受信装置。

【請求項13】 上記雑音レベル検出手段は、上記受信部の音声レベルが所定値以上のときに、上記送話用マイクロフォンに入力される音声レベルを検出することを特徴とする請求項10、11又は12記載の音声信号送受信装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

本発明は、音声信号送受信装置に関し、特に、デジタル信号処理により音声

信号を高能率で圧縮する音声信号送受信装置に関する。

【0002】

【従来の技術】

近年、低ビットレート、すなわち、4.8 ～9.6 Kbpsでの音声符号化方法には、ベクトル和励起リニア予測 (VSELP: Vector Sum Excited Linear Prediction) 等のコード励起リニア予測 (CELP: Code Excited Linear Prediction) が提唱されている。

【0003】

このVSELPについての技術内容は、モトローラ・インコーポレーテッドによる特表平2-502135号公報の「改良されたベクトル励起源を有するデジタル音声コーダ」及び「VECTOR SUM EXCITED LINEAR PREDICTION (VSELP): SPEECH CODING AT 8 KBPS: Ira A. Gerson and Jasiuk: Paper presented at the Int. Conf. on Acoustics, Speech and Signal Processing -April 1990」に記載されている。

【0004】

上記VSELPを用い音声を高能率で圧縮するデジタル信号処理による音声符号化装置としては、VSELPエンコーダがある。このVSELPエンコーダは、入力された音声信号から音声のフレームパワー、反射係数及び線形予測係数、ピッチ周波数、コードブック、ピッチ及びコードブックのゲイン等のパラメータを分析し、この分析パラメータを用いて、音声を符号化している。このような音声を高能率で圧縮するデジタル信号処理による音声符号化装置である上記VSELPエンコーダは、携帯用電話装置に適用されている。

【0005】

この携帯用電話装置は、室外で使用する事が多いので、周囲の背景雑音により、しばしば通話が聞きずらくなることがある。これは、雑音によるマスキング効果によって受話者の最少可聴値が上昇し、受話音声の明瞭度や了解度が劣化するためである。これに対して、送話側では雑音の抑圧又は話者の声量のアップ、受話側では再生音量のアップ、又全体として話者と電話器の密接な音響カップリング等が必要になる。このため、携帯用電話装置には回りの環境に応じて受話音量

を手動で切り換えるスイッチがある。

【0006】

【発明が解決しようとする課題】

ところで、上述したように携帯用電話装置を使用する際、回りの環境に応じて受話音量を手動で切り換えるのは面倒である。この受話音量の切り換えを自動で行えるようになれば便利である。

【0007】

この受話音量の切り換えを自動で行おうとする場合、回りの雑音レベルを正確に検出できるか否かが問題となってくる。通話中に入力音声用（送話用）のマイクロフォンから混入する雑音には様々なものがあるがこれらの雑音（以下背景雑音という）と音声の区間を分離するのはかなり難しかった。

【0008】

これに対し、信号に含まれている基本周期やピッチ等を検出したり、信号波形のゼロクロスの頻度を見たり、周波数成分の分布をみることを併用することで、背景雑音区間と音声区間とを区別することが考えられてきた。これらの手法は、簡便である反面、誤検出がしばしばあった。又、精度を向上するためのアルゴリズムも提案されているが、かなりの演算量を要する。例えば長時間の平均的な線形予測符号化（LPC）係数を用いて、入力信号に逆フィルタリングを施し、その残差レベルをモニタする方法なども提案されているが演算量が多くなってしまっていた。

【0009】

本発明は、上記実情に鑑みてなされたものであり、少ない演算量ながら高精度、高信頼度で背景雑音を検出し、該検出背景雑音に応じて再生音量である受話音量をコントロールして明瞭度の高い受話音を聞かせる音声信号送受信装置の提供を目的とする。

【0010】

【課題を解決するための手段】

本発明に係る音声信号送受信装置は、上記課題を解決するために、デジタル信号処理により音声信号を高効率で圧縮する音声信号送信用符号化回路を有する

音声信号送受信装置において、上記音声信号送信用符号化回路で得られる分析パラメータを用いて雑音区間を検出する雑音区間検出手段と、上記雑音区間検出手段で検出された雑音区間の雑音レベルを検出する雑音レベル検出手段と、上記雑音レベル検出手段で検出された雑音レベルに応じて受話音量を制御する制御手段とを有することを特徴としている。

【0011】

ここで、上記雑音区間検出手段は、入力音声信号の1フレームにつき上記分析パラメータとして1次の線形予測符号化係数を用い、該1次の線形予測符号化係数が所定のしきい値よりも小さいときには該1フレームを雑音区間としてもよい。

【0012】

また、上記雑音区間検出手段は、入力音声信号の1フレームにつき上記分析パラメータとしてピッチ成分の強弱を示すピッチゲインを用い、該ピッチゲインが所定の範囲内にあるときに該1フレームを雑音区間としてもよい。

【0013】

また、上記雑音区間検出手段は、上記分析パラメータとしてピッチ周波数を用い、入力音声信号の1フレームの該ピッチ周波数成分が0であるときに該1フレームを雑音区間としてもよい。

【0014】

また、上記雑音区間検出手段は、上記分析パラメータとしてフレームパワーを用い、入力音声信号の1フレームの該フレームパワーが所定のしきい値よりも小さいときに該1フレームを雑音区間としてもよい。

【0015】

また、上記雑音区間検出手段は、現在のフレームと過去のフレームでのフレームパワーの変化量が所定のしきい値を越えたときには、現在のフレームを雑音区間としていても、該現在のフレームを音声区間とする。

【0016】

また、上記雑音区間検出手段は、複数連続フレームの上記分析パラメータの値を考慮して、雑音区間の検出を行ってもよい。



【0017】

また、上記雑音レベル検出手段は、上記雑音区間検出手段で検出された雑音区間の雑音レベル出力に対してフィルタ処理を施すのが好ましい。

【0018】

また、上記雑音レベル検出手段が雑音レベル出力に対して施すフィルタ処理は、最小値フィルタ処理であってもよいし、メディアンフィルタリング或いは平滑化フィルタ処理でもよい。

【0019】

また、本発明に係る音声信号送受信装置は、上記課題を解決するために、送信部及び受信部を備える音声信号送受信装置において、上記送信部に送話音声入力が無いときに送話用マイクロフォンに入力される音声レベルを雑音レベルとして検出する雑音レベル検出手段と、上記雑音レベル検出手段で検出された雑音レベルに応じて受話音量を制御する制御手段とを有することを特徴とする。

【0020】

ここで、上記雑音レベル検出手段は、上記送信部の上記送話用マイクロフォンに送信通話用電源オン直後に入力される音声レベルを検出する。

【0021】

また、上記雑音レベル検出手段は、上記送信部の着信信号の待機状態において、所定期間毎に上記送信用マイクロフォンに入力される音声レベルを検出してもよい。

【0022】

また、上記雑音レベル検出手段は、上記受信部の音声レベルが所定値以上のときに、上記送話用マイクロフォンに入力される音声レベルを検出してもよい。

【0023】

【作用】

本発明に係る音声信号送受信装置は、雑音区間検出手段が音声信号送信用符号化回路で得られる分析パラメータを用いて雑音区間を検出するので、少ない演算量ながら高精度、高信頼度の下に雑音区間を検出でき、雑音レベル検出手段がその雑音区間の雑音レベルを検出し、制御部がその雑音レベルに応じて再生音声の

音量を制御するので、明瞭度の高い受話音を供給できる。

【0024】

また、本発明に係る音声信号送受信装置は、雑音レベル検出手段が送信部に送話音声入力がないときに、送話用マイクロフォンに入力される音声レベルを雑音レベルとして検出し、制御手段が該検出された雑音レベルに応じて受話音量を制御するので、背景雑音の影響に左右されない明瞭度の高い受話音を供給できる。

【0025】

【実施例】

以下、本発明に係る音声信号送受信装置の好ましい実施例について、図面を参照しながら説明する。

図1は、本発明の実施例となる携帯電話装置の概略構成を示すブロック回路図である。

【0026】

この実施例は、図1に示すように、デジタル信号処理により音声信号を高効率で圧縮するベクトル和励起リニア予測（VSELP: Vector Sum Excited Linear Prediction）エンコーダ3と、このVSELPエンコーダ3で得られる分析パラメータを用いて背景雑音（以下適宜に雑音という）区間を検出する雑音区間検出回路4と、この雑音区間検出回路4で検出された雑音区間の雑音レベルを検出する雑音レベル検出回路5と、この雑音レベル検出回路5で検出された雑音レベルに応じて受話音量を制御するマイクロコンピュータ6とを有して構成されている。

【0027】

上記VSELPエンコーダ3を用いた音声符号化方法としては、アナリシスバイシンセシス（Analysis by synthesis）によるコードブックサーチにより、低ビットレートによる高品質音声伝送を実現している。また、VSELPを用いた音声符号化方法を適用した音声符号化装置（音声コーダ）においては、入力音声信号の特性を形成するピッチ等をコードブックに記憶されたコードベクトルを選択することで励起させて音声を符号化している。この符号化の際に用いるピッチ周波数等のパラメータには、フレームパワー、反射係数及び線形予測係数、コード

ブック、ピッチ及びコードブックのゲイン等がある。

【0028】

本実施例は、これらの分析パラメータの内、フレームパワー $R_0$ 、ピッチ成分の強弱を示すピッチゲイン $P_0$ 、1次の線形予測符号化係数 $\alpha_1$ 及びピッチ周波数に関するラグLAGを背景雑音検出に利用する。例えばフレームパワー $R_0$ を利用するのは、音声レベルと雑音レベルが同じになることはほとんどないためであり、ピッチゲイン $P_0$ を利用するのは、周囲雑音がほぼランダムであるとすれば、この周囲雑音はピッチをほとんど持たないと考えられるためである。

【0029】

また、1次の線形予測符号化係数 $\alpha_1$ を用いるのは、この $\alpha_1$ が大か小かで、周波数の高域成分が強いかわるいは低域成分が強いかを判定できるからである。通常、背景雑音は、周波数の高域成分に集中しており、上記1次の線形予測符号化係数 $\alpha_1$ から背景雑音を検出できる。この1次の線形予測符号化係数 $\alpha_1$ は、直接型の高次のFIRフィルタを2次のFIRフィルタのカスケードに分解したときの逆関数 $Z^{-1}$ の係数の和である。したがって、零点が $0 < \theta < \pi/2$ の範囲にある時、1次の線形予測符号化係数 $\alpha_1$ は大きくなる。よって、この $\alpha_1$ が所定のしきい値より大きいときは、低域にエネルギーの集中した信号ということになり、所定のしきい値より小さいときは、高域にエネルギーの集中した信号ということになる。

【0030】

ここで、 $\theta$ と周波数との関係について説明しておく。

サンプリング周波数を $f$ とすると、 $0 \sim f/2$ の周波数がデジタルフィルタ等のデジタルシステムにおいて、 $0 \sim \pi$ に相当する。例えば、サンプリング周波数 $f$ を8 KHzとすると、 $(0 \sim 4 \text{ KHz})$ は $(0 \sim \pi)$ に相当し、よって、 $\pi/2 = 2 \text{ KHz}$ となる。したがって、 $\theta$ が小さいほど周波数成分が低域になる。また、 $\theta$ が小さくなれば、 $\alpha_1$ は大きくなるので、 $\alpha_1$ と所定のしきい値との関係を調べることで低域成分が強いのか高域成分が強いのか分かる。

【0031】

次に、上記雑音区間検出回路4は、上記VSELPエンコーダ3から上記分析

パラメータすなわちフレームパワー $R_0$ 、ピッチ成分の強弱を示すピッチゲイン $P_0$ 、1次の線形予測符号化係数 $\alpha_1$ 及びピッチ周波数に関するラグLAGを受け取り、雑音区間を検出する。これは、携帯電話装置が小型化されていく現在、ディジタル信号処理(DSP)装置やメモリの大きさが制限されており、演算量を増やすのを避けるためにも有効である。

#### 【0032】

上記雑音レベル検出回路5は、上記雑音区間検出回路4で検出された雑音区間の音声レベルすなわち送話用音声レベルを検出する。ここで、検出される送話用音声レベルは、上記雑音区間検出回路4の上記分析パラメータを用いた判定により最終的に雑音区間とされたフレームのフレームパワー $R_0$ の値としてもよい。但し、検出ミスの可能性があるので、このフレームパワー $R_0$ を後述するように例えば5タップの最小値フィルタ等に入力する。

#### 【0033】

上記マイクロコンピュータ6は、上記雑音区間検出回路4での雑音区間検出と上記雑音レベル検出回路5での雑音レベル検出のタイミングを制御すると共に、該雑音レベルに応じて再生音声の音量を制御する。

#### 【0034】

このような本実施例は、以下に説明するように全体的に構成されている。

すなわち、送話用マイクロフォン1で電気信号とされた入力音声信号は、アナログ/ディジタル(A/D)変換器2によりディジタル信号とされて、VSELPエンコーダ3に供給される。このVSELPエンコーダ3は、ディジタル信号とされた入力信号を分析し、情報圧縮をし、符号化を行う。この際、入力音声信号のフレームパワー、反射係数及び線形予測係数、ピッチ周波数、コードブック、ピッチ及びコードブックのゲイン等の分析パラメータを用いている。

#### 【0035】

上記VSELPエンコーダ3で情報圧縮、符号化が施されたデータは、ベースバンド信号処理回路7に供給され、同期信号の付加、フレーミング、誤り訂正符号等を付加される。そして、ベースバンド信号処理回路7からの出力データは、RF送受信回路8に供給され、必要な周波数に変調されてアンテナ9から送信さ

れる。

【0036】

ここで、上記VSELPエンコーダ3が用いた分析パラメータの内、上述したようにフレームパワー $R_0$ 、ピッチ成分の強弱を示すピッチゲイン $P_0$ 、1次の線形予測符号化係数 $\alpha_1$ 及びピッチ周波数に関するラグLAGは、上記雑音区間検出回路4に供給される。この雑音区間検出回路4は、上記フレームパワー $R_0$ 、ピッチ成分の強弱を示すピッチゲイン $P_0$ 、1次の線形予測符号化係数 $\alpha_1$ 及びピッチ周波数に関するラグLAGを用いて、雑音区間の検出を行う。この雑音区間検出回路4で最終的に雑音区間であるとされたフレームに関する情報（フラグ情報）は、上記雑音レベル検出回路5に供給される。

【0037】

上記雑音レベル検出回路5には、上記A/D変換器2からのデジタル入力信号も供給されており、上記フラグ情報に応じて雑音区間の信号レベルを検出する。この場合の信号レベルは、上述したようにフレームパワー $R_0$ としてもよい。

【0038】

上記雑音レベル検出回路5で検出された雑音レベルデータは、制御部であるマイクロコンピュータ6に供給される。このマイクロコンピュータ6には、後述するように受信側レベル検出回路11からの情報も供給されており、これらの情報を下に後述するように可変利得アンプ13の利得を可変することにより受話音量を制御する。

【0039】

この受話音量とは、本実施例の携帯電話装置に送信されてきた通話相手からの信号を再生するときの音量である。この通話相手からの信号は、アンテナ9により受信され、RF送受信回路8に供給される。

【0040】

このRF送受信回路8によりベースバンドに復調された相手側からの入力音声信号は、ベースバンド信号処理回路7に供給され、所定の信号処理が施される。このベースバンド信号処理回路7からの信号は、VSELPデコーダ10に供給される。このVSELPデコーダ10は、この情報を元に音声信号をデコードす

る。デコードされた音声信号は、ディジタル／アナログ（D／A）変換器12に供給され、アナログ音声信号に変換される。

【0041】

上記VSELPデコーダ10からのデコードされた音声信号は、受信側レベル検出回路11にも供給される。この受信側レベル検出回路11は、受信側音声のレベルの検出を行い、現在受話音声（相手側からの入力音声信号）があるか否かを判定する。この受信側レベル検出回路11からの検出情報は、上記マイクロコンピュータ6に供給される。

【0042】

上記D／A変換器12からのアナログ音声信号は、可変利得アンプ13に供給される。この可変利得アンプ13の利得は、上述したように上記マイクロコンピュータ6により可変されているので、スピーカ14から発せられる再生音量（受話音量）は、雑音（背景雑音）に応じて、マイクロコンピュータ6により制御される。

【0043】

なお、このマイクロコンピュータ6には、表示装置15、電源回路16及びキーボード17が接続されている。表示装置15は、この本実施例である携帯電話装置が使用可能であるか、キーボード17で使用者が押圧したキースイッチが何であるか等を表示する。

【0044】

次に、本実施例を構成する上記雑音レベル検出回路5での雑音レベルの検出について以下に説明する。

先ず、雑音レベルを検出する区間は、上記雑音区間検出回路4で検出された雑音区間であることが条件となる。この雑音区間を検出するタイミングは、上述したように上記マイクロコンピュータ6で制御される。この雑音区間の検出は、上記雑音レベル検出回路5での雑音レベルの検出を補助するためのものである。すなわち該当するフレームが有声音である音声かあるいは雑音であるかを判定し、雑音であるという判定であれば雑音レベルの検出が可能となる。当然のことながら、より精度の高い雑音レベルの検出は、雑音のみが存在する時に行うのが良い

のは明らかである。したがって、本実施例では、送話音声入力が無いときに送話用マイクロフォン1に入力される音声レベルを送話用音声レベル検出手段でもある雑音レベル検出回路5に検出させている。

【0045】

まず、雑音レベルの初期値として例えば使用者が設定した音量レベルに対して-20dBを設定する。この初期設定値に対して後述するように検出された雑音レベルが大きいと判断された時には、受信部での再生音量レベルを上昇させる。

【0046】

雑音レベルは、フレーム毎の入力音声为背景雑音区間であれば、上述したように検出しやすい。このため、本実施例では、送信部の送信通話用電源がオンとされた直後、送信部の着信信号の待機状態及び通話中であって受信部の音声レベルが所定値以上のときに入力される音声を背景雑音とし、この間のフレームの雑音レベルを検出している。

【0047】

ここで、送信部の送信通話電源がオンとされることは、使用者が本実施例の携帯電話装置の使用を開始する意思表示である。このとき、本実施例は、通常、内部の各回路の自己診断を行い、次に、使用者がアンテナ9を張ると基地局との接続を確認した上でスタンバイ状態に入る。これらの一連の動作を経て初めて使用者からの入力（入力音声）を受けるので、使用者がこの間に音声を発することはない。したがって、この一連の動作の最中に上記送話用マイクロフォン1を使用して音声レベルを検出すれば、検出された音声レベルは周囲のノイズレベルすなわち背景雑音レベルである。なお、同様に、通話開始直前で使用者が発振操作をした最中又は直後も背景雑音レベルの検出が可能である。

【0048】

また、送信部の着信信号の待機状態とは、受話部の電源をオンにして、相手側からの通話信号の着信を待ち受けている状態である。この状態のときには、当然のことながら通話中ではないので、使用者の送話音声が無いと考えられる。そこで、この待ち受け状態に、送話用マイクロフォン1を用いて周囲の音量レベルを測定すれば、背景雑音レベルを検出できる。なお、この測定は、適当な間隔で行

い平均化してもよい。

【0049】

以上により、送信部の送信通話電源がオンとされた直後及び送信部の着信信号の待機状態で背景雑音レベルが推定でき、それに応じた音声処理によって通話がスタートできるが、その後の背景雑音レベルの変化に対しては、通話中もダイナミックに追従することが好ましい。そこで、本実施例では、通話中での受信部の音声レベルに応じて背景雑音レベルの検出を行っている。

【0050】

この通話中での受信部の音声レベルに応じた雑音レベルの検出は、上述したように受話側のVSELPエンコーダ3で用いられる分析パラメータにより雑音区間を検出してから行うのが好ましい。

【0051】

例えば、フレームパワー $R_0$ をモニタしそのレベルがある基準のレベル以上であるときや、相手が話しているときを利用して雑音レベルを検出すること等により、より確実に雑音の検出ができるので、相手が話しているときの再生音量をリアルタイムで制御でき、より快適な通話品質が実現できる。

【0052】

このように本実施例では、送信部の送信通話用電源がオンとされた直後、送信部の着信信号の待機状態及び通話中であって送信部の音声がないときに、上記マイクロコンピュータ6が上記雑音区間検出回路4及び上記雑音レベル検出回路5の検出タイミングを制御している。

【0053】

次に、上記雑音区間検出回路4での雑音区間検出動作について、図2及び図3に示すフローチャートを参照しながら説明する。

【0054】

まず、図2のフローチャートが開始されるとステップS1では、上記VSELPエンコーダ3からフレームパワー $R_0$ 、ピッチ成分の強弱を示すピッチゲイン $P_0$ 、1次の線形予測符号化係数 $\alpha_1$ 及びピッチ周波数に関するラグLAGを受け取る。



## 【0055】

本実施例においては、上記ステップS1で供給された各分析パラメータを用いた以下の各ステップでの判別を基本的に3フレームで行うことにした。これは、1フレームだけで背景雑音の判別を行うと誤りが多くなるためである。そして、3フレームに渡り各パラメータの範囲を見ながら、雑音区間を判別したら、ノイズフラグを“1”とし、そうでなければ“0”にセットする。3フレームの内訳は、現在のフレームと1、2フレーム前までのフレームである。

## 【0056】

このような連続した3フレームを通しての分析パラメータによる判別を以下の各ステップで行う。

## 【0057】

まず、ステップS2では、入力音声のフレームパワー $R_0$ が3フレーム連続して所定のしきい値 $R_{0th}$ より小さいか否かを判別する。ここで、YES ( $R_0$ が3フレーム連続して $R_{0th}$ より小さい)と判別するとステップS3に進み、NO ( $R_0$ が3フレーム連続して $R_{0th}$ 以上である)と判別するとステップS9に進む。この所定のしきい値 $R_{0th}$ は、それ以上のレベルをノイズではなく、音声と見なす値である。すなわち、このステップS2は、信号レベルのチェックである。

## 【0058】

ステップS3では、入力音声の1次の線形予測符号化(LPC)係数 $\alpha_1$ が3フレーム連続して所定のしきい値 $\alpha_{th}$ より小さいか否かを判別する。ここでYES ( $\alpha_1$ が3フレーム連続して $\alpha_{th}$ より小さい)と判別するとステップS4に進み、NO ( $\alpha_1$ が3フレーム連続して $\alpha_{th}$ 以上である)と判別するとステップS9に進む。この所定のしきい値 $\alpha_{th}$ は、雑音进行分析したときにはほとんど表れることのない値になっている。すなわち、このステップS3は、音声スペクトルの傾きのチェックである。

## 【0059】

ステップS4では、現在の入力音声のフレームのフレームパワー $R_0$ の値が“5”より小さいか否かを判別する。ここで、YES ( $R_0$ が5より小さい)と判

別すると、ステップS5に進み、NO ( $R_0$  が5以上である)と判別すると、ステップS6に進む。ここで、“5”をしきい値としたのは、フレームパワー $R_0$ が“5”より大である場合のフレームは、有声音である確率が高いためである。

【0060】

ステップS5では、入力音声信号のピッチゲイン $P_0$ の値が3フレーム連続して0.9より小さく、かつ現在のピッチゲイン $P_0$ が0.7より大きいかな否かを判別する。ここで、YES (ピッチゲイン $P_0$ の値が3フレーム連続して0.9より小さく、かつ現在のピッチゲイン $P_0$ が0.7より大きい)と判別すると、ステップ8に進み、NO (ピッチゲイン $P_0$ の値が3フレーム連続して0.9以上、また現在のピッチゲイン $P_0$ が0.7以下である)と判別すると、ステップS9に進む。上記ステップS3から上記ステップS5までは、ピッチ成分の強弱のチェックである。

【0061】

ステップS6では、上記ステップS4での判別結果(NO:  $R_0$  が5以上である)を受けて、そのフレームパワー $R_0$ が5以上20未満であるかな否かを判別する。ここでYES ( $R_0$  が5以上20未満である)と判別するとステップS7に進み、NO ( $R_0$  が5以上20未満でない)と判別するとステップS9に進む。

【0062】

ステップS7では、入力音声信号のピッチゲイン $P_0$ の値が3フレーム連続して0.85より小さく、かつ現在のピッチゲイン $P_0$ が0.65より大きいかな否かを判別する。ここで、YES (ピッチゲイン $P_0$ の値が3フレーム連続して0.85より小さく、かつ現在のピッチゲイン $P_0$ が0.65より大きい)と判別すると、ステップ8に進み、NO (ピッチゲイン $P_0$ の値が3フレーム連続して0.85以上、また現在のピッチゲイン $P_0$ が0.65以下である)と判別すると、ステップS9に進む。

【0063】

ステップS8では、上記ステップS5又は上記ステップS7でのYESの判別結果を受けて、ノイズフラグを“1”とする。ノイズフラグを“1”とすることは、そのフレームを雑音とすることである。

【0064】

ステップS9では、上記ステップS2、上記ステップS3、上記ステップS5、上記ステップS6及び上記ステップS7での判別がNOとされた場合に、ノイズフラグを“0”とし、該当フレームを音声であるとする。

【0065】

次に、図3のフローチャートに移る。

ステップS10では、入力音声信号のピッチラグLAGが0であるか否かの判別を行う。ここでYES（LAGが0である）と判別すると、ピッチ周波数を表すLAGが0の場合は、音声である確率はほとんどないので、そのフレームを雑音とする。すなわち、ステップS11に進みノイズフラグを“1”とする。ここでNO（LAGが0でない）と判別するとステップS12に進む。

【0066】

ステップS12では、フレームパワー $R_0$ が2以下であるか否かを判別する。ここで、YES（ $R_0$ が2以下である）と判別するとステップS13に進み、NO（ $R_0$ が2より大きい）と判別するとステップS14に進む。このステップS12は、フレームパワー $R_0$ がかなり小さいか否かを判別しており、YESと判定すると次のステップS13でノイズフラグを“1”とし、そのフレームを雑音としている。

【0067】

ステップS13では、上記ステップS11と同様にそのフレームを雑音とすべく、ノイズフラグを“1”とする。

【0068】

ステップS14では、現在のフレームのフレームパワー $R_0$ から1つ前のフレームパワー $R_0$ を減算し、その絶対値が3を超えるか否かを判別する。現在のフレームと1つ前のフレームでのフレームパワー $R_0$ の変化が急に大きくなる時には、そのフレームを音声フレームとするためである。すなわち、このステップS14でYES（現在のフレームと1つ前のフレームのフレームパワー $R_0$ の変化が急激に大きくなった）と判定するとステップS16に進み、ノイズフラグを“0”とし、そのフレームを音声フレームとする。また、ここで、NO（現在の

フレームと1つ前のフレームのフレームパワー $R_0$ の変化が急激に大きくならない)と判別すると、ステップS15に進む。

【0069】

ステップS15では、現在のフレームのフレームパワー $R_0$ から2つ前のフレームパワー $R_0$ を減算し、その絶対値が3を超えるか否かを判別する。現在のフレームと2つ前のフレームでのフレームパワー $R_0$ の変化が急に大きくなる時には、そのフレームを音声フレームとするためである。すなわち、このステップS15でYES(現在のフレームと2つ前のフレームのフレームパワー $R_0$ の変化が急激に大きくなった)と判定するとステップS16に進み、ノイズフラグを“0”とし、そのフレームを音声フレームとする。また、ここで、NO(現在のフレームと2つ前のフレームのフレームパワー $R_0$ の変化が急激に大きくならない)と判別すると、ステップS17に進む。

【0070】

ステップS17では、最終的にノイズフラグを“0”又は“1”と決定し、そのフラグ情報を上記雑音レベル検出回路5に供給する。

【0071】

以上、図2及び図3に示したフローチャートによる雑音区間検出回路4での動作により得られたフラグ情報に応じて上記雑音レベル検出回路5は、雑音区間の音声レベルを検出する。

【0072】

ところで、上記雑音区間検出回路4での雑音区間検出では、完全に音声区間と雑音区間とを区別することは出来ず、また、音声を誤って雑音として検出してしまうことが起こりえる。この検出誤りは、ほとんどが音声の子音部で起きる。背景雑音が子音部と略々同じくらいのレベルで混入している場合は、誤検出しても報告される雑音レベルが変わらないので問題ないが、そうでない場合、特に雑音がほとんど混入していないような場合では、レベルが場合によっては、20~30dBも違うため、かなり問題になってくる。そこで、本実施例では、誤検出した場合でもそのまま検出した雑音区間の雑音レベルを用いるのではなく、平滑化などにより誤検出の影響が少なくなるようにした。

## 【0073】

このような平滑化等の手段により誤検出の影響を少なくした雑音レベルの検出について、図4を参照しながら説明する。

## 【0074】

図4において、入力端子20には、上記A/D変換器2からのデジタル入力信号が供給される。また、入力端子21には、上記雑音区間検出回路4からのフラグ情報がデジタルシグナルプロセッサ(DSP)で構成される雑音レベル検出回路5の雑音レベル決定部5aに入力されるように供給される。この雑音レベル決定部5aには、入力端子22からのフレームパワー $R_0$ も供給されている。すなわち、この雑音レベル決定部5aは、雑音区間検出回路4からのフラグ情報又はフレームパワー $R_0$ を基に入力音声信号の雑音レベルを決定している。具体的には、図3に示したフローチャートのステップS17において、最終的にノイズフラグが“1”とされたときのフレームパワー $R_0$ の値を背景雑音レベルと見なしている。

## 【0075】

このとき、検出ミスの可能性があるので、この $R_0$ の値を例えば5タップの最小値フィルタ5bに入力する。この $R_0$ は、背景雑音と認められた時のみ入力する。この最小値フィルタ5bの出力は、マイクロコンピュータ6等の制御用CPUに適当な周期(例えば100msec毎)で入力する。ここで、最小値フィルタ5bの出力が更新されていないときは、前の値を繰り返し使用する。この最小値フィルタ5bは、後述するメディアンフィルタのようにタップ中の真ん中の値を出力するものではなく、最小値を出力する。同じタップ数の場合、連続した4フレームまでの検出誤りに対応できる。また、それ以上の誤りについても、最小値を報告レベルとするため、影響をなるべく少なくできる。

## 【0076】

上記マイクロコンピュータ6では、入力された信号レベル $R_0$ の信頼度をより向上するために、該信号レベル $R_0$ を更に5タップのメディアンフィルタ6aに入力させている。このメディアンフィルタ6aは、検出誤りが続いてもレベルの報告を誤りにくいようにする。このフィルタリングは、フィルタのタップ中の値

を小さい順に並べ変え、その中の中間値を出力するものである。5タップのメディアンフィルタは、連続した2フレームまでは検出誤りがおきても、報告レベルを間違えることはない。

【0077】

上記メディアンフィルタ6aの出力信号は、ボリューム位置調整部6bに供給される。このボリューム位置調整部6bは、上記メディアンフィルタ6aの出力信号を基に上記可変利得アンプ13の利得を可変する。このようにして、上記マイコン6は、再生音量である受話音量を制御する。具体的には、使用者の設定したボリューム位置を中心（基点）として、音量の増減をコントロールするものである。また、使用者がボリュームを調節した直前の雑音レベルを記憶しておき、そのレベルと現在の背景雑音レベルの変化分に基づき出力音量を増減してもよい。

【0078】

なお、ここで、用いられるフィルタとしては、検出した背景雑音レベルの平滑化を行う1次のローパスフィルタ等の平滑化フィルタでもよい。ローパスフィルタの度合いによっては、検出を誤ってレベルが急に変化しても追従が遅くなるためレベル差を小さくできる。

【0079】

このようにすれば、雑音レベルを誤検出した場合でも、誤検出の影響を少なくできる。

【0080】

ここで、検出された雑音レベルによる受話音量制御の方法を説明しておく。

受話音量をコントロールする際、通常は、上述したように初期設定された音量を背景雑音に応じて変化させる。もし、使用者が音量ボリュームを手動で変えた場合は、その音量を基に背景雑音のレベルに応じて受話音量をコントロールするようにする。

【0081】

具体的には、初期値として例えば図5に示すように5段階（1～5：小から大へ変化する）の雑音レベルに応じた受話音量レベル（a、b、c、d、e）が与

えられていて、この値に基づいてコントロールされる。

【0082】

例えば、使用者が手動調整できる音量ボリュームつまみを上げた場合（例えば（-1）、（0）、（+1）の3段階があり（0）から（+1）に上げた場合）、雑音レベルが上がる（例えばレベル1であった雑音は、 $1 + (+1)$  となりレベルは2となる）。ここで、本発明では、雑音レベルに応じて受話音量レベルも上昇させている（例えば雑音レベルが2であれば、受話音量レベルはbとなる。

【0083】

また例えば、使用者が手動調整できる音量ボリュームつまみを下げた場合（例えば（-1）、（0）、（+1）の3段階があり（0）から（-1）に下げた場合）、雑音レベルが下がる（例えばレベル3であった雑音は、 $3 + (-1)$  となりレベルは2となる）。ここで、本発明では、雑音レベルに応じて受話音量レベルも下降させている（例えば雑音レベルが2であれば、受話音量レベルはbとなる。

【0084】

つまり、使用者が手動調整できる音量ボリュームつまみを上下した場合、音量ボリュームつまみを変更する直前の雑音レベルと受話音量の対応を覚えておき、使用者が音量ボリュームつまみを変えた時点で雑音レベルと音量レベルの対応（マッピング）を変えることで、ダイナミックに受話音量の基準値を変更する。このようにすることで、話者の意図した（変更した）音量（音量ボリュームつまみにより手動調整された音量）を基に雑音レベルに応じた受話音量のコントロールができる。

【0085】

次に、以上に述べた本実施例の有効性を確かめるために、シミュレーションによって実際に背景雑音検出を行った例を述べておく。

通常、室内騒音の基準としてH o t hスペクトルで表されているものが一般に使われているが、室外で 사용되는ことの多い携帯電話装置にこのH o t hスペクトルを適用するのは難しい。そこで、シミュレーションには実際に室外で録音してきた雑音を使用した。この雑音は、2つの駅（これをA駅、B駅とする）構

内で録音した。そして、音声と雑音をコンピュータ上でデジタル波形で加算した場合、雑音を試験室内で流しその状態で携帯電話装置を使ってマイクを通して話したときの音声を録音した場合、雑音のない音声の場合の3通りについて検討を行った。なお、雑音の混入レベルは、70dBspl 程度の騒音環境を想定した。

【0086】

このシミュレーションとしては、固定少数点によるシミュレーションを行い、検出の頻度や誤り及び検出した雑音レベルなどについて調べた。

【0087】

背景雑音を検出した例を図6乃至図9に示す。これら図6乃至図9は、サンプルとしてA駅又はB駅構内で録音した背景雑音を流しながら携帯電話装置を使って話したときの音声と、検出した背景雑音の結果である。

【0088】

図6がA駅構内で録音した背景雑音を流しながら男声で「人間は豊かな自然を求めています」と話したときの結果であり、図7がA駅構内で録音した背景雑音を流しながら女声で「健康のため無理をしないで下さい」と話したときの結果である。また、図8がB駅構内で録音した背景雑音を流しながら男声で「人間は豊かな自然を求めています」と話したときの結果であり、図9がB駅構内で録音した背景雑音を流しながら女声で「健康のため無理をしないで下さい」と話したときの結果である。

【0089】

それぞれの検出結果において、図中矩形部分が背景雑音と思われる部分を検出した区間である。音声部分と雑音部分を完全に分離することは出来ないが、数10ms単位での検出が出来ており、また音声部分を誤って検出してしまうことがほとんどない。子音部での背景雑音検出誤りは、前述した平滑化等の手段を用いることで報告レベルの間違いを避けることができた。特に、最小値フィルタリングを使うことによりほとんど検出誤りによるレベル報告の間違いを避けることができた。

【0090】

なお、このような雑音検出のシミュレーションとしては、上記固定少数点によ



るシミュレーションだけではなく、例えばワークステーション上で浮動少数点によるシミュレーションを行ってもよく、得られる検出結果はほとんど同じである。

#### 【0091】

以上より、本実施例の携帯電話装置は、VSELPエンコーダで用いられている分析パラメータを使用して雑音区間検出を行うので、少ない演算量ながら高精度、高信頼度で背景雑音を検出でき、該背景雑音に応じて再生音量をコントロールするので、明瞭度の高い受話音を提供できる。

#### 【0092】

なお、本発明に係る音声信号送受信装置は、上記実施例にのみ限定されるものではなく、例えば雑音区間の検出には分析パラメータを1つだけ用いることも可能である。さらに、複数の連続したフレームを考慮するのではなく、1フレームのみで検出することも可能である。しかし、これらの場合、その精度は、上記本実施例よりも落ちてしまうことは否めない。さらに、雑音区間の検出の流れも上記フローチャートに示したものに限定されるものでないことはいうまでもない。

#### 【0093】

##### 【発明の効果】

本発明に係る音声信号送受信装置は、雑音区間検出手段が音声信号送信用符号化回路で得られる分析パラメータを用いて雑音区間を検出するので、少ない演算量ながら高精度、高信頼度の下に背景雑音を検出でき、雑音レベル検出手段がその雑音区間の雑音レベルを検出し、制御部がその雑音レベルに応じて受話音量を制御するので、信頼度の高い検出となり、その検出された背景雑音のレベル情報に基づき受話音量の制御などが容易にかつ確実に行え、明瞭度の高い受話音を供給できる。

#### 【0094】

また、本発明に係る音声信号送受信装置は、雑音レベル検出手段が送信部に送話音声入力がないときに、送話用マイクロフォンに入力される音声レベルを雑音レベルとして検出し、制御手段が該検出された音声レベルに応じて受話音量を制御するので、背景雑音の影響に左右されない明瞭度の高い受話音を供給できる。

【0095】

また、受話音量を上げると人間は話す音量を上げる傾向にあるので、送信音の明瞭度を上げることもできる。

【図面の簡単な説明】

【図1】

本発明に係る音声送受信装置の実施例の回路構成を説明するためのブロック回路図である。

【図2】

図1に示した実施例の背景雑音検出回路の動作を説明するためのフローチャートである。

【図3】

図1に示した実施例の背景雑音検出回路の動作を説明するためのフローチャートである。

【図4】

背景雑音レベルを誤差の影響から防ぐための手段を説明するための図である。

【図5】

本実施例により、検出された雑音レベルによる受話音量制御の具体例を説明するための図である。

【図6】

固定少数点によるシミュレーションを行って得られた背景雑音検出結果を示す図である。（A駅構内での雑音を流して男声で話した場合）

【図7】

固定少数点によるシミュレーションを行って得られた背景雑音検出結果を示す図である。（A駅構内での雑音を流して女声で話した場合）

【図8】

固定少数点によるシミュレーションを行って得られた背景雑音検出結果を示す図である。（B駅構内での雑音を流して男声で話した場合）

【図9】

固定少数点によるシミュレーションを行って得られた背景雑音検出結果を示す

図である。(B駅構内での雑音を流して女性で話した場合)

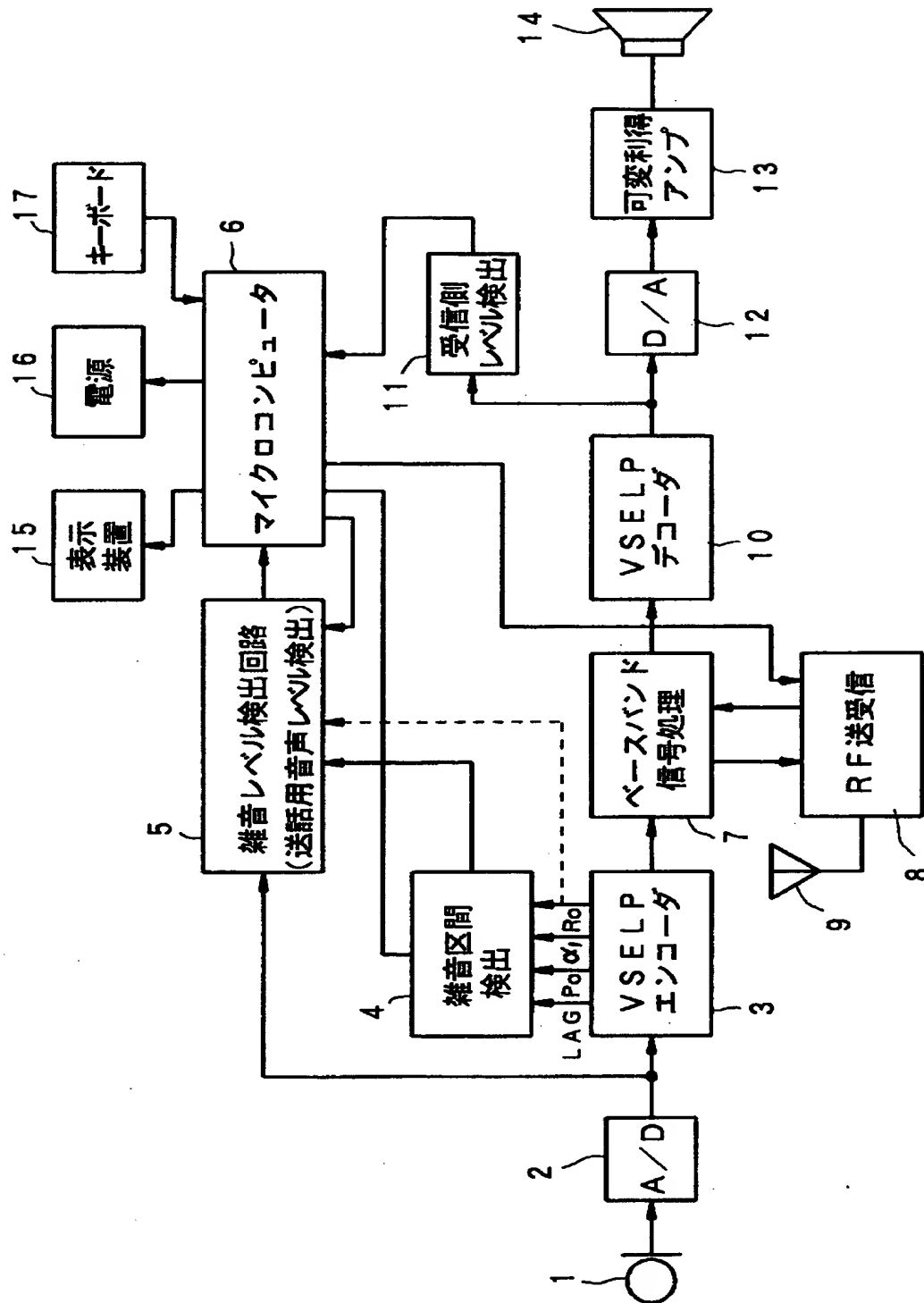
【符号の説明】

- 1 . . . . . 送話用マイクロフォン
- 2 . . . . . アナログ／デジタル (A／D) 変換器
- 3 . . . . . V S E L P エンコーダ
- 4 . . . . . 雑音区間検出回路
- 5 . . . . . 雑音レベル検出回路
- 6 . . . . . マイクロコンピュータ
- 7 . . . . . ベースバンド信号処理回路
- 8 . . . . . R F 送受信回路
- 9 . . . . . アンテナ
- 10 . . . . . V S E L P デコーダ
- 11 . . . . . 受信側レベル検出回路
- 12 . . . . . デジタル／アナログ (D／A) 変換器
- 13 . . . . . 可変利得アンプ
- 14 . . . . . スピーカ

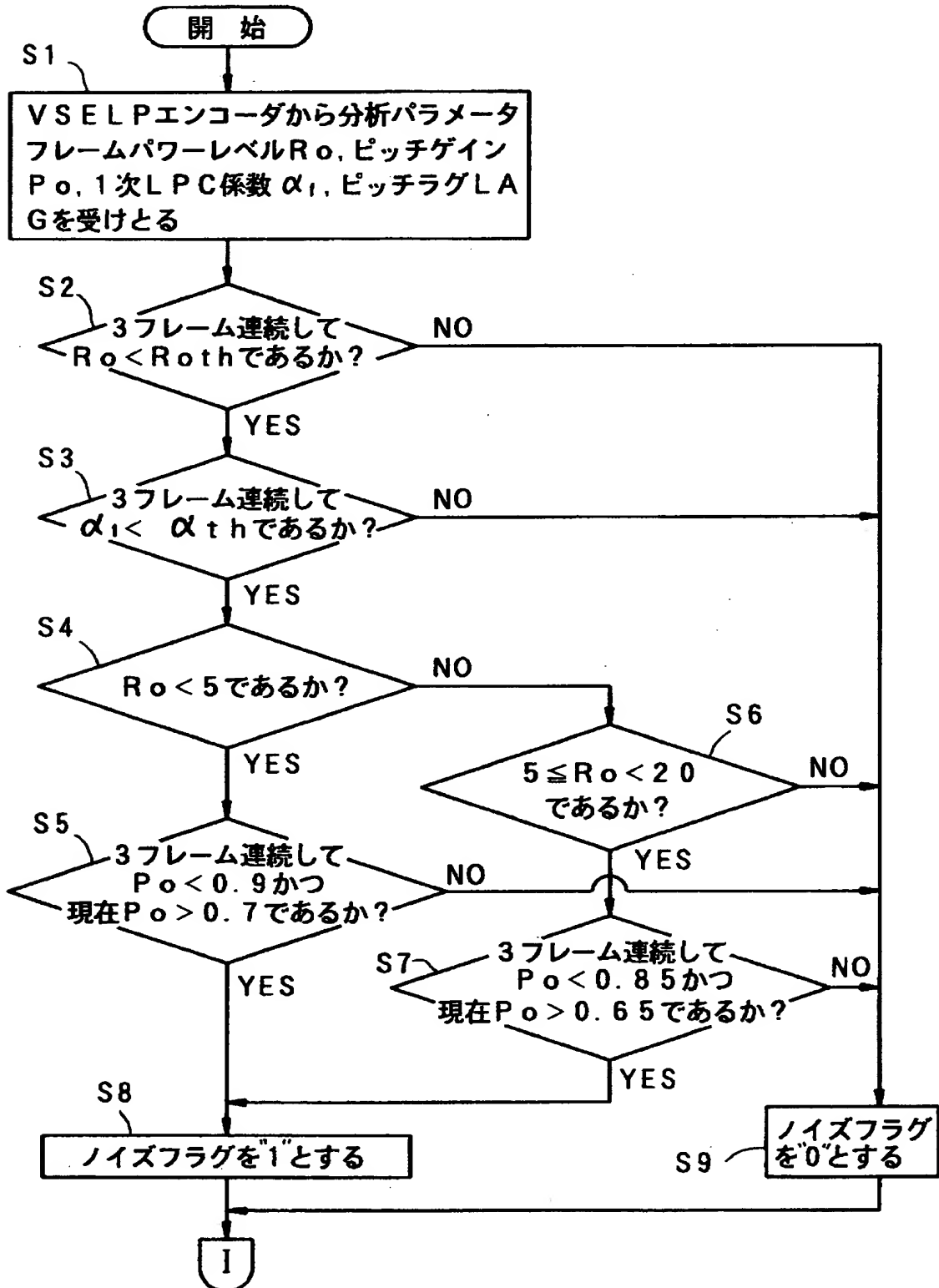
【書類名】

図面

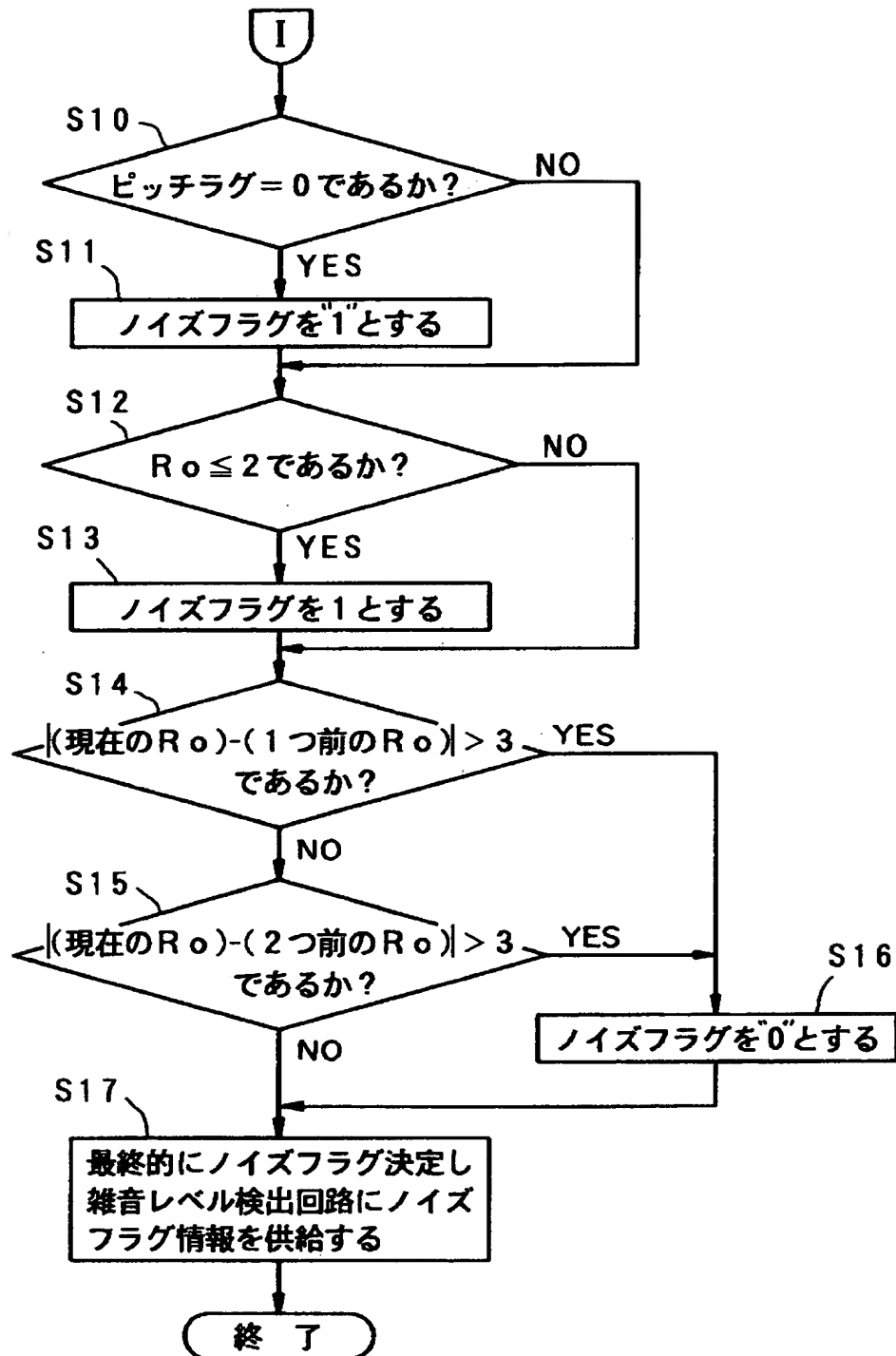
【図1】



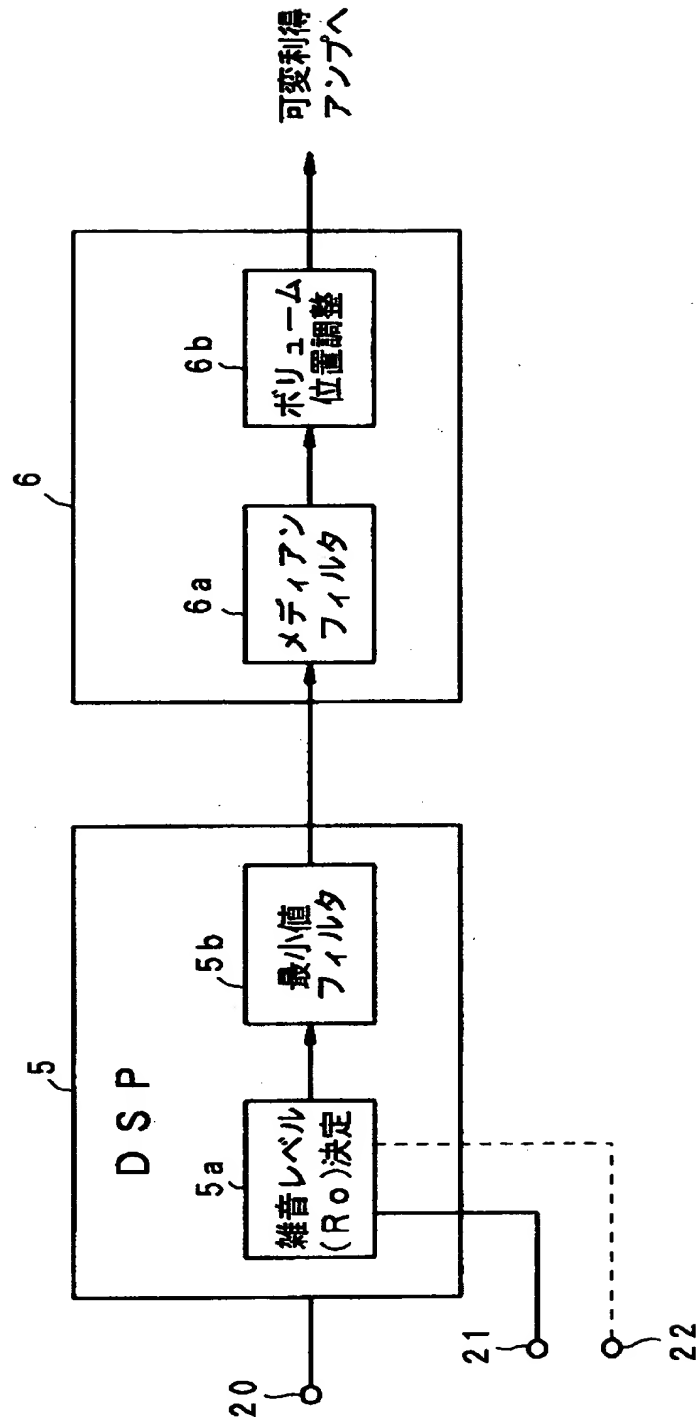
【図2】



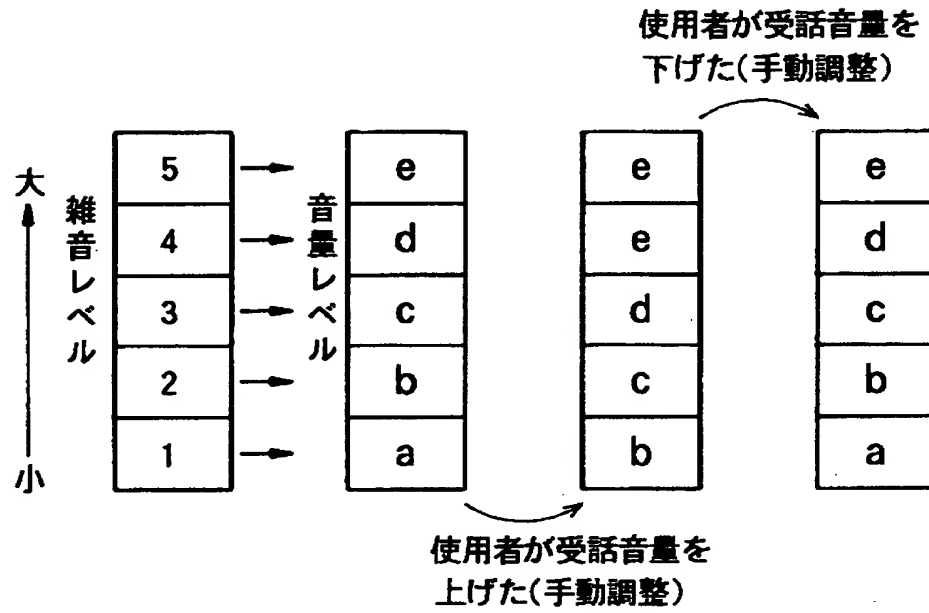
【図3】



【図4】

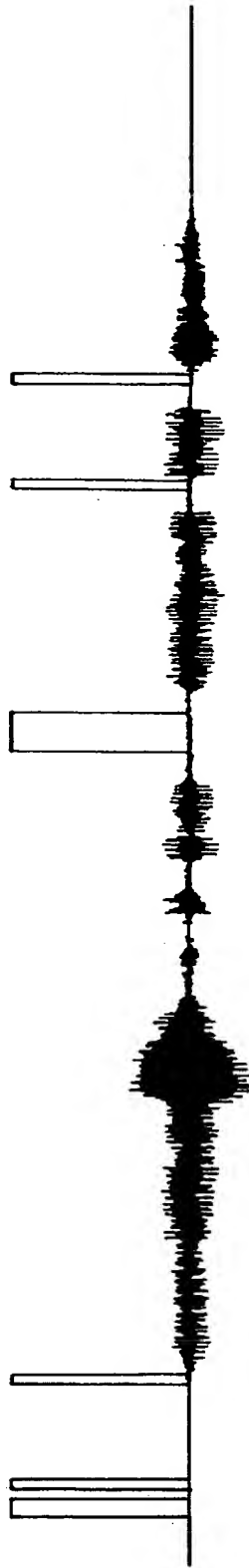


【図5】



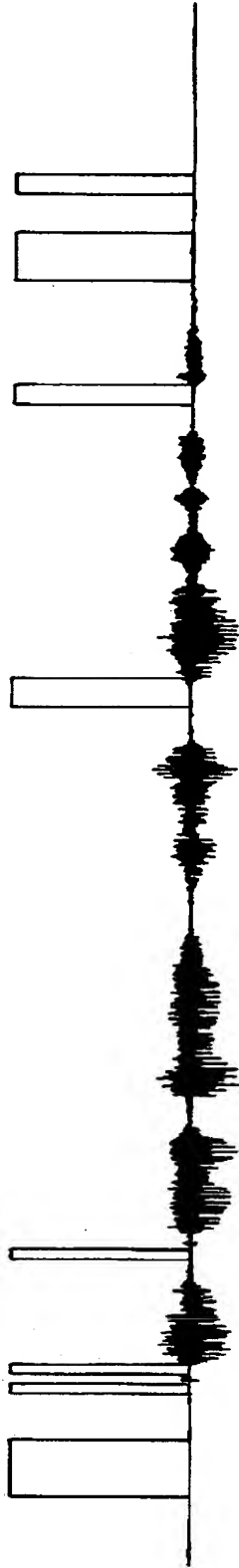


【図6】



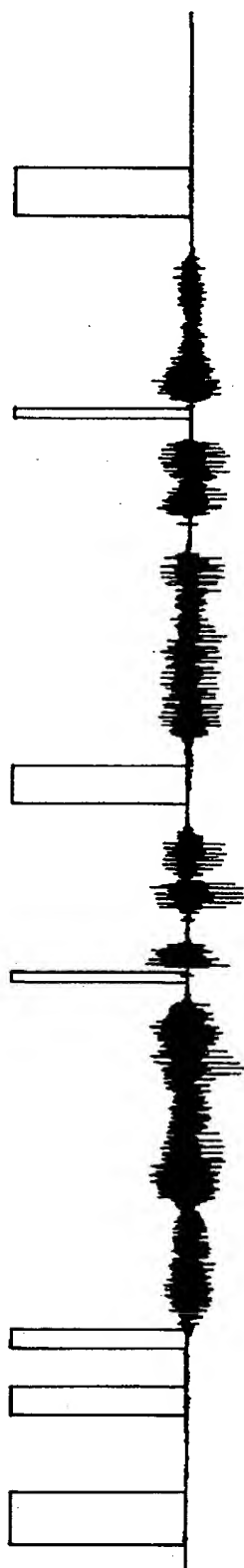
特平 5-182138

【図7】



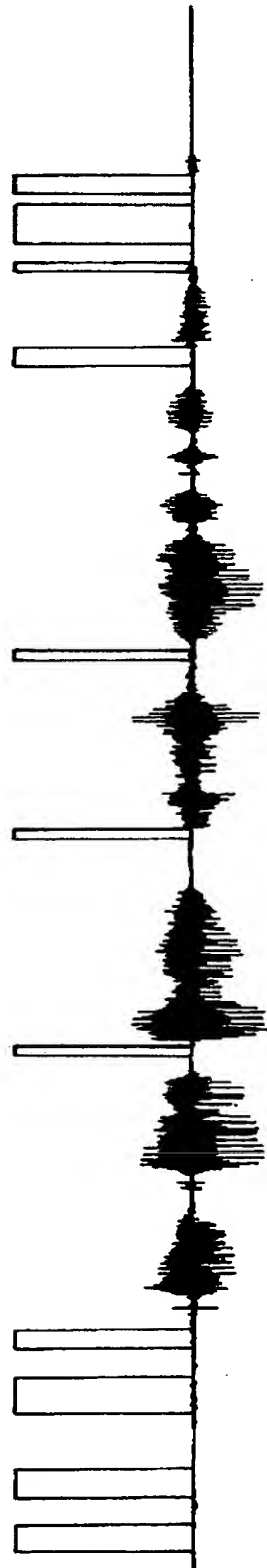
特平 5-182138

【図8】



特平 5-182138

【図9】



【書類名】 要約書

【要約】

【構成】 V S E L P エンコーダ 3 は、ディジタル信号処理により高能率符号化された音声信号を符号化する。雑音区間検出回路 4 は、上記 V S E L P エンコーダ 3 で得られる分析パラメータを用いて背景雑音区間を検出する。雑音レベル検出回路 5 は、上記雑音区間検出回路 4 で検出された雑音区間の雑音レベルを検出するので、信頼度の高い背景雑音レベルを検出でき、マイクロコンピュータ 6 は、上記雑音レベル検出回路 5 で検出された上記背景雑音レベルに応じて受話音量を制御する。

【効果】 受話音量の制御が容易に行え、明瞭度の高い受話音を供給できる。

【選択図】 図 1



【書類名】  
【訂正書類】

職権訂正データ  
特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000002185

【住所又は居所】 東京都品川区北品川6丁目7番35号

【氏名又は名称】 ソニー株式会社

【代理人】 申請人

【識別番号】 100067736

【住所又は居所】 東京都港区虎ノ門2-6-4 第11森ビル 小池  
国際特許事務所

【氏名又は名称】 小池 晃

【代理人】 申請人

【識別番号】 100086335

【住所又は居所】 東京都港区虎ノ門2丁目6番4号 第11森ビル  
小池国際特許事務所

【氏名又は名称】 田村 榮一

【代理人】 申請人

【識別番号】 100096677

【住所又は居所】 東京都港区虎ノ門二丁目6番4号 第11森ビル  
小池国際特許事務所

【氏名又は名称】 伊賀 誠司

出 願 人 履 歴 情 報

識別番号 [000002185]

1. 変更年月日 1990年 8月30日  
[変更理由] 新規登録  
住 所 東京都品川区北品川6丁目7番35号  
氏 名 ソニー株式会社